

Cyanogen Bromide (CNBr) Cleavage

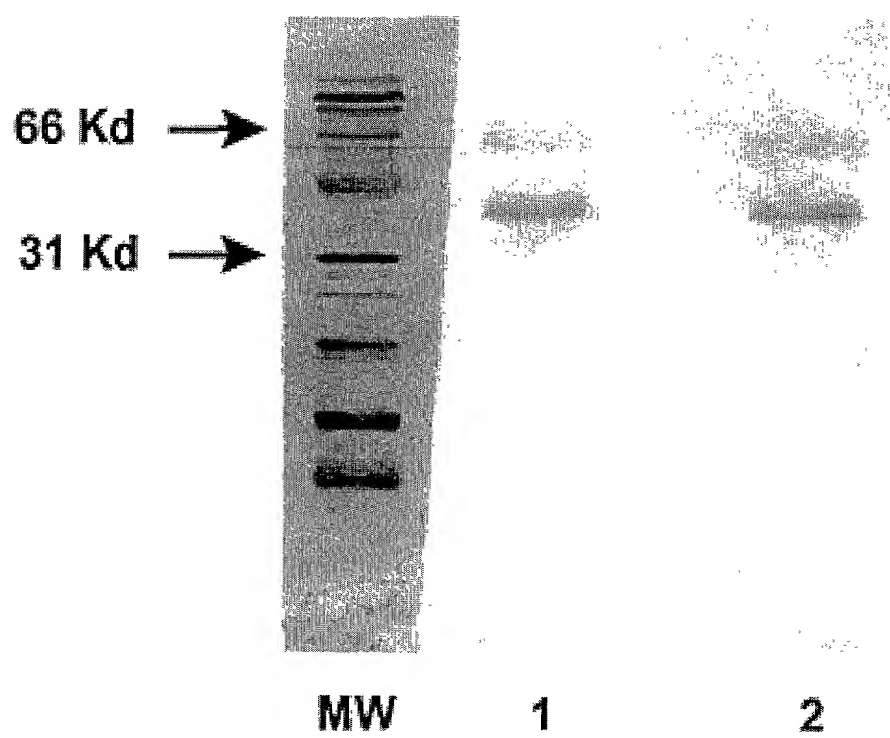
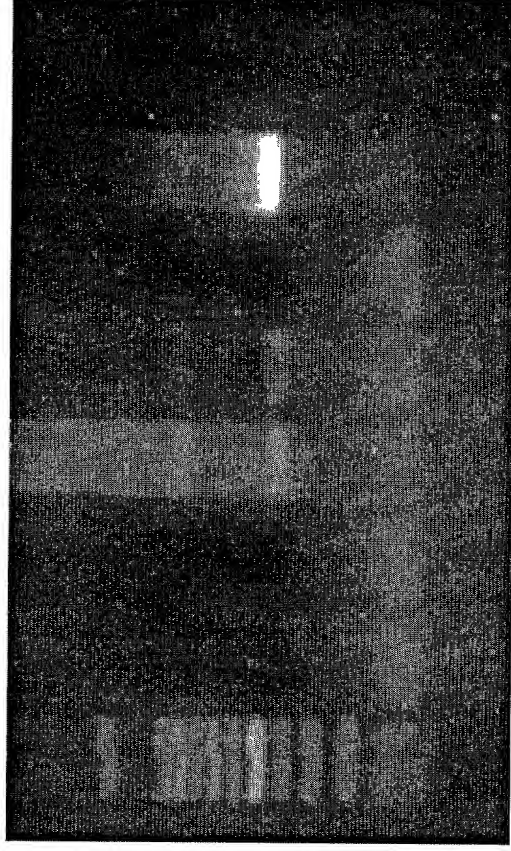


Figure 1

A

1 2 3 4 5 6



500 bases

B

1 2 3 4 5 6 7 8 9 10 11 12



CA125



β-Tubulin

Figure 2

A Strategy for Placing Repeat Sequences in Contiguous Order Using Overlap Sequence Alignment

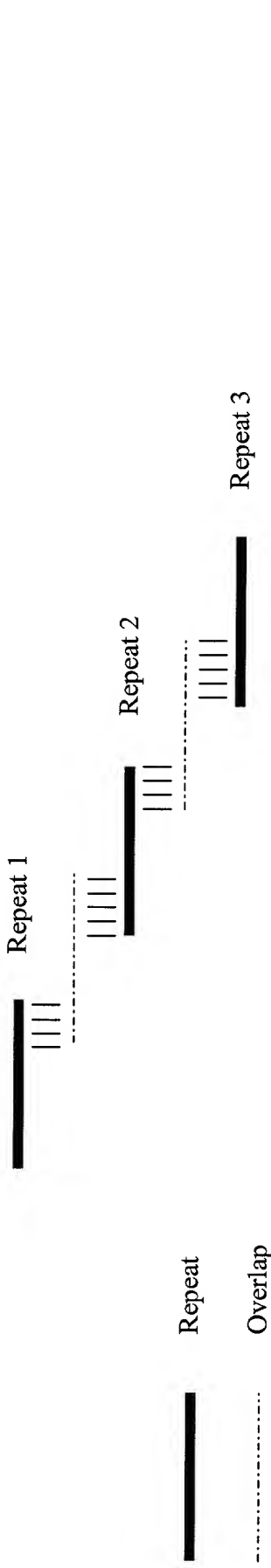


Figure 3 (SEQ ID NOS: 158, 159, 160, and 161)

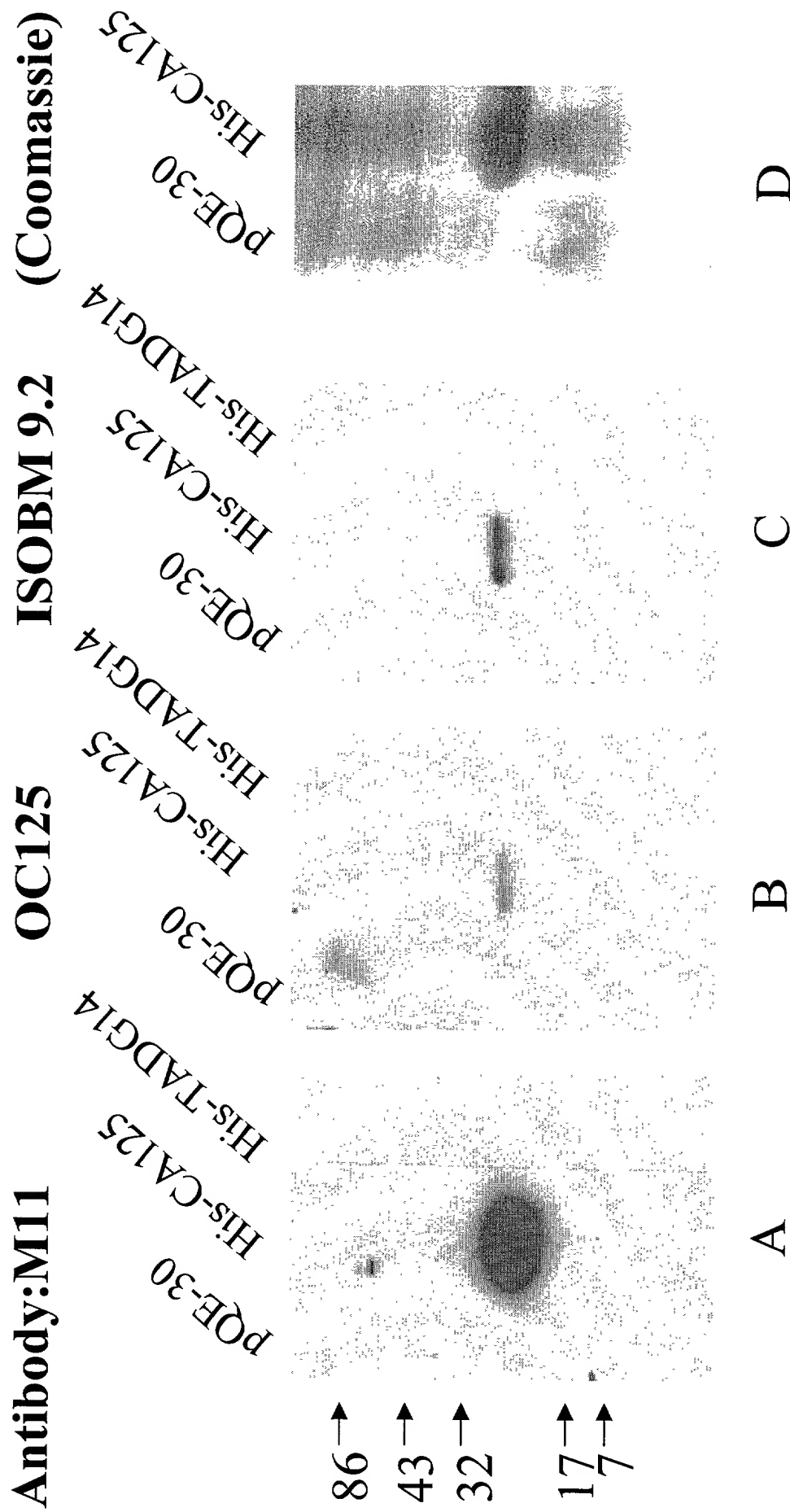


Figure 4

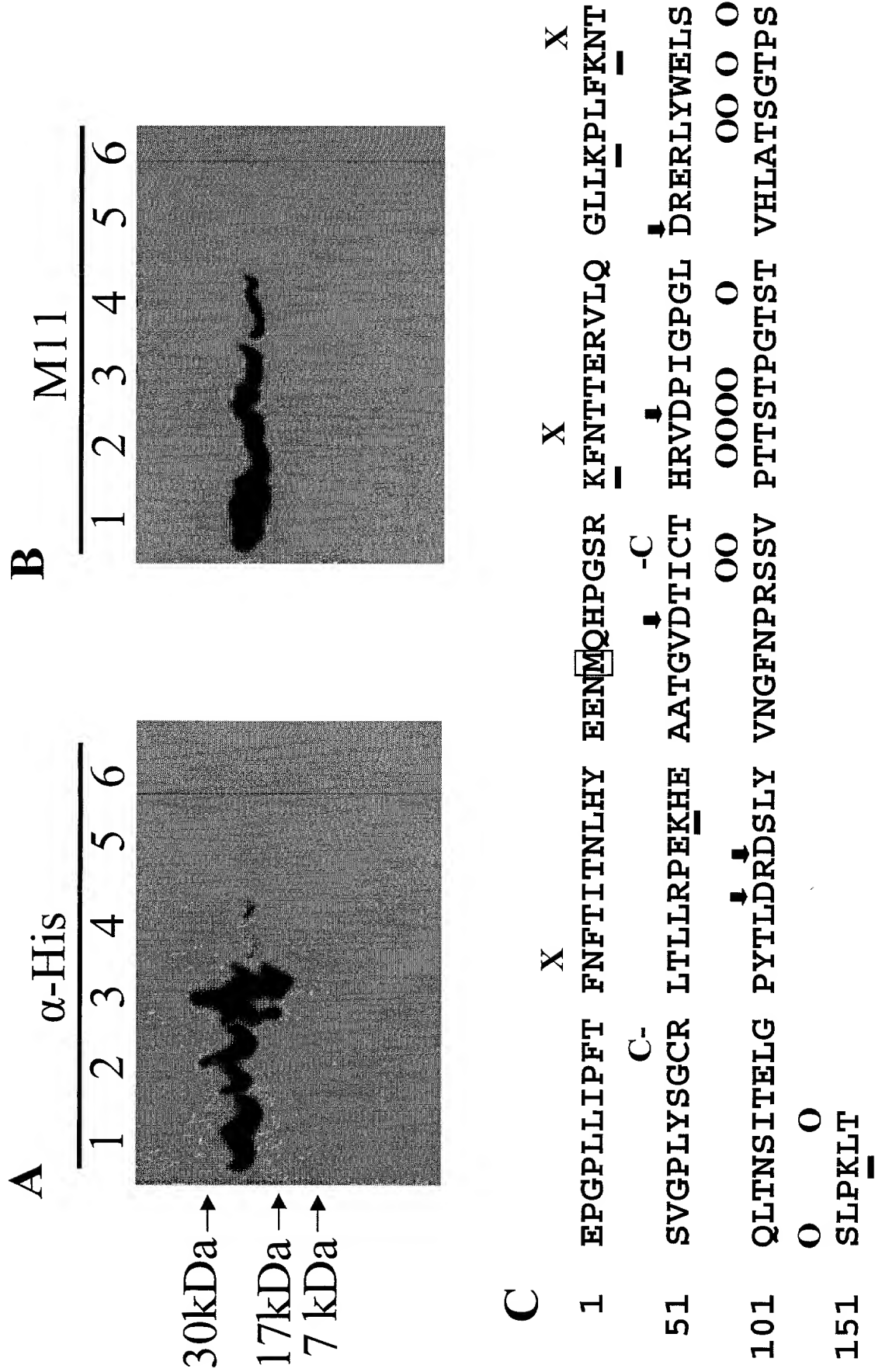


Figure 5 (SEQ ID NO: 150)

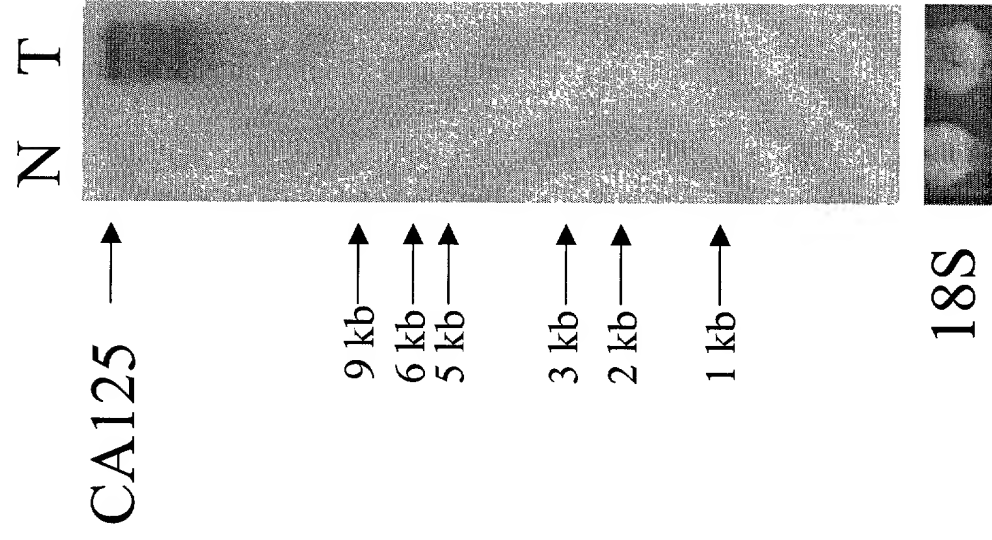


Figure 6

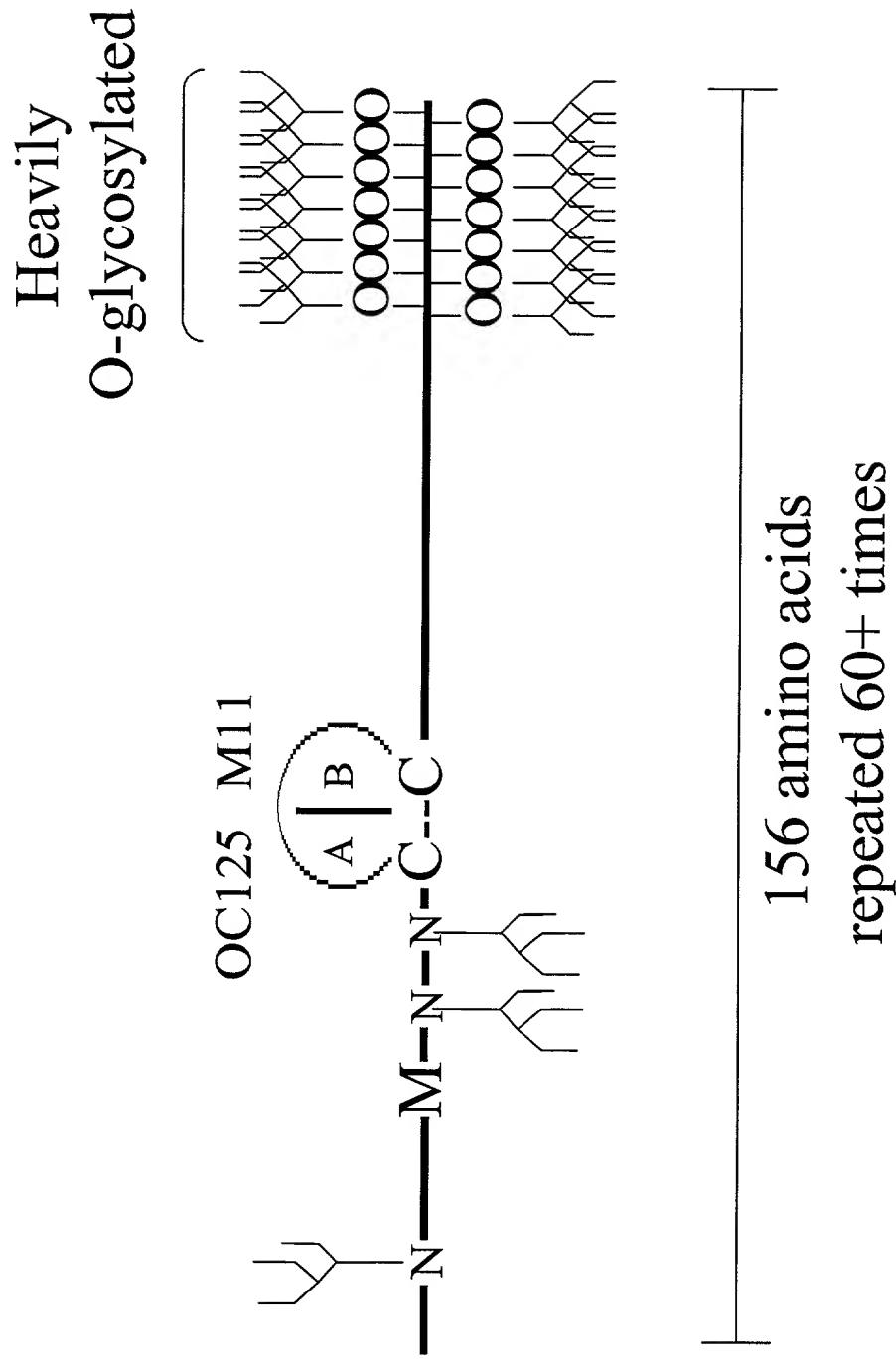


Figure 7A

1042503" 36259663

Genomic Structure of a 156 Amino Acid Repeat Sequence of CA125

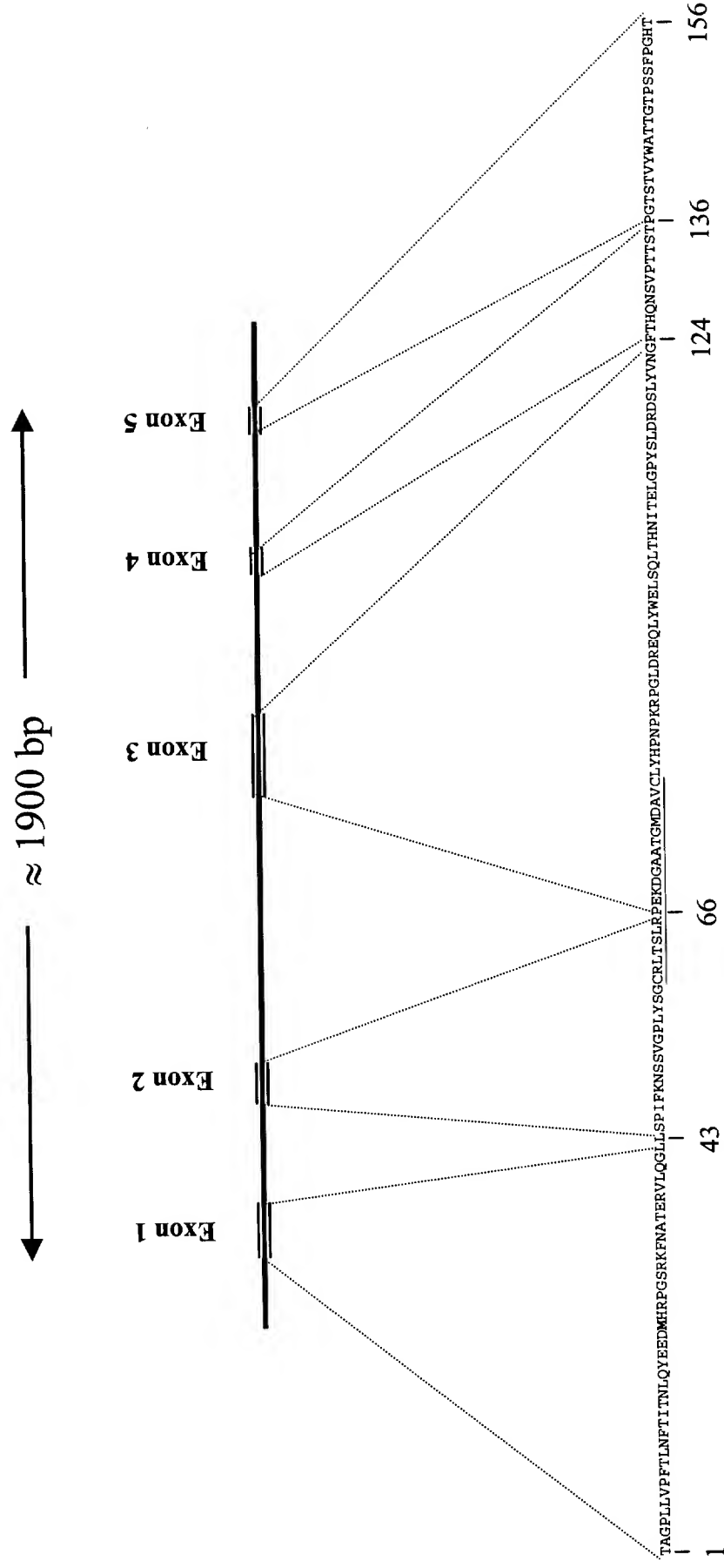


Figure 7B (SEQ ID NO: 163)

Exon 1

1

42

ATVPFMVPFTLNFTTITNLQYEEDMRHPGSRKFNATERLQGL (SEQ ID NO: 164)
TAVPLLVPFTLNFTTITNLQYGEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 165)
VPGPLLVPFTLNFTTITNLQYEEAMRHPGSRKFNTTERVLQGL (SEQ ID NO: 166)
APGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFSTTERVLQGL (SEQ ID NO: 167)
APGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 168)
APGPLLVPFTLNFTTITNLQYEVDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 169)
SAGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 170)
AAGPLLMPFTLNFTTITNLQYEEDMRRTGSRKFNTMESVLQGL (SEQ ID NO: 171)
TASPLLVLFTINCTITNLQYEEDMRRTGSRKFNTMESVLQGL (SEQ ID NO: 172)
AAGPLLVPFTLNFTTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 173)
TAGPLLIPFTLNFTTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 174)
TAGPLLVPFTLNFTTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 175)
TAGPLLVLFTLNFTTITNLKYEEDMRHPGSRKFNTTERVLQTL (SEQ ID NO: 176)
TAGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFNATERVLQGL (SEQ ID NO: 177)
TAGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 178)
TAGPLLVPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 179)
APVPLLIPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 180)
ATGPVLLPFTLNFTTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 181)
AAGPLLVPFTLNFTTITNLQYEEDMHPGSRKFNTTERVLQGL (SEQ ID NO: 182)
SAGPLLVPFTLNFTTITNLQYEEDMHPGSRKFNTTERVLQGL (SEQ ID NO: 183)
TASPLLVLFTINFTITNQRYEENMHPGSRKFNTTERVLQGL (SEQ ID NO: 184)
TASPLLVLFTINFTITNLRYEENMHPGSRKFNTTERVLQGL (SEQ ID NO: 185)
EPGPLLIPFTFNFTTITNLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 186)
EPGPLLIPFTFNFTTITNLRYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 187)
APVPLLIPFTLNFTTITNLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 188)
APVPLLIPFTLNFTITDLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 189)
AASPLLVLFTLNFTTITNLRYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 190)
TAGPLLVPFTLNFTTITNLKYEEDMHCPGSRKFNTTERVLQSL (SEQ ID NO: 191)
AASHLLILFTLNFTTITNLRYEENMW.PGSRKFNTTERVLQGL (SEQ ID NO: 192)
TGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNTDNVMKHL (SEQ ID NO: 193)
AMGYHLKTLTLNFTISNLQYSPDMGKGSATFNSTEGVLQHLL (SEQ ID NO: 194)

Figure 7C

Exon 2

43 65

LKPLFRNSSLEYLYSGCRLASLR (SEQ ID NO: 195)
LKPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 196)
LKPLFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 197)
LKPLFKSTSVGPLYSGCRLTLLR (SEQ ID NO: 198)
LKPLFKSTSVGPLYSSCRLTLLR (SEQ ID NO: 199)
LKPLFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 200)
LGPIFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 201)
LGPMFKNTSVGLLYSGCRLTLLR (SEQ ID NO: 202)
LGPMFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 203)
LGPMFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 204)
LGPLFKNSSVGPLYSGCRLISLR (SEQ ID NO: 205)
LGPLFKNSSVDPLYSGCRLTSLR (SEQ ID NO: 206)
LSPIFKNSSVGPLYSGCRLTSLR (SEQ ID NO: 207)
LSPIFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 208)
LSPLFQRSSLGARYTGCRVIALR (SEQ ID NO: 209)
LRPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 210)
LRPLFKNTSVGPLYSGSRLTLLR (SEQ ID NO: 211)
LRPLFKNTSIGPLYSSCRLTLLR (SEQ ID NO: 212)
LRPLFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 213)
LRPVFKNTSVGLLYSGCRLTLLR (SEQ ID NO: 214)
LRPVFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 215)
LRSLFKSTSVGPLYSGCRLTLLR (SEQ ID NO: 216)
LRSLFKSTSVGPLYSGCRLTSLR (SEQ ID NO: 217)
LTPLFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 218)
LTPLFRNTSVSSLYSGCRLTLLR (SEQ ID NO: 219)
LMPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 220)
RPLFQKSSM.GPFYLGQCQLISLR (SEQ ID NO: 221)

Figure 7C

Exon 3

66

123

PEKDSSAMAVDAICTHRPDPEDLGLDRERLYWELSNLTNGIQELGPYTLDRNSLYVNG (SEQ ID NO: 222)
PEKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNG (SEQ ID NO: 223)
PKKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNG (SEQ ID NO: 224)
PEKDGATATGVDAICTHHPDPKSPRLDREQLYWELSQLTHNITELGHYALDNDSLFVNG (SEQ ID NO: 225)
PEKDGATATGVDAICTHRPDPGGLDREQLYLELSQLTHSITELGPYTLDRDLSLYVNG (SEQ ID NO: 226)
PEKDGATATGMDAVCLYHPNPKRPGGLDREQLYWELSQLTHNITELGPYSLDRDLSLYVNG (SEQ ID NO: 227)
PEKDGATATGMDAVCLYHPNPKRPGGLDREQLYCELSQLTHNITELGPYSLDRDLSLYVNG (SEQ ID NO: 228)
PEKDGATATRVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRVSLYVNG (SEQ ID NO: 229)
PKKDGAATKVDICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDLSLYVNG (SEQ ID NO: 230)
PKKDGAATKVDICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDLSLYVNG (SEQ ID NO: 231)
PEKDGATATRVDAICTHRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRHSLYVNG (SEQ ID NO: 232)
PEKDGATATRVDAICTHRPDKPIGLDRQQLYWELSQLTHSITELGPYTLDRDLSLYVNG (SEQ ID NO: 233)
SEKDGAATGVDAICTHHLDPKSPGLNREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 234)
SEKDGAATGVDAICTHRLDPKSPGLDREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 235)
SEKDGAATGVDAICTHRLDPKSPGVDRQQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 236)
SEKDGAATGVDAICTHRVDPKSPGVDRQQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 237)
SEKDGAATGVDAICTHHLNPQSPGLDREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 238)
PEKGAATGVDTICTHRLDPLNPGLDREQLYWELSKLTRGIIELGPYLLDRGSLYVNG (SEQ ID NO: 239)
PEKGAATGVDAICTHRLDPKSPGLNREQLYWELSQLTHGIKELGPYTLDRNSLYVNG (SEQ ID NO: 240)
PEKGAATGVDAICTHRLDPKSPGLDREQLYWELSQLTHGIKELGPYTLDRNSLYVNG (SEQ ID NO: 241)
PEKHGAATGVDAICTLRLDPTGPGGLDREQLYWELSQLTNSVTELGPYTLDRDLSLYVNG (SEQ ID NO: 242)
PEKHGAATGVDAICTLRLDPTGPGGLDREQLYWELSQLTNSITELGPYTLDRDLSLYVNG (SEQ ID NO: 243)
PEKHEAATGVDTICTHRVDPGPGGLDREQLYWELSQLTNSITELGPYTLDRDLSLYVNG (SEQ ID NO: 244)
PEKQEAATGVDTICTHRVDPGPGGLDREQLYWELSQLTNSITELGPYTLDRDLSLYVNG (SEQ ID NO: 245)
PEKQEAATGVDTICTHRVDPGPGGLDREQLYWELSQLTNSITELGPYTLDRDLSLYVDG (SEQ ID NO: 246)
PEKKAATRVDAICTHHPDPQSPGLNREQLYWELSQLTHGITELGPYTLDRDLSLYVDG (SEQ ID NO: 247)
SVKGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQLTHGITRLGPYSLDKDLSLYVNG (SEQ ID NO: 248)
PEKDGAATGVDTTCTYHPDPVGPGLDIQQLYWELSQLTHGVTQLGFYVLDRLDLSLYVNG (SEQ ID NO: 249)

Figure 7C

Exon 4

124 135
FTHRSMPTTST (SEQ ID NO: 250)
FTHRSMPTTST (SEQ ID NO: 251)
FTHRSTSVPTSST (SEQ ID NO: 252)
FTHRSTSVPTTST (SEQ ID NO: 253)
FTHRSSVPTTSS (SEQ ID NO: 254)
FTHRSSVSTTST (SEQ ID NO: 255)
FTHRSSVAPTST (SEQ ID NO: 256)
FTHRSSGLTTST (SEQ ID NO: 257)
FTHRSTFGLTTST (SEQ ID NO: 258)
FTHRSSFLTST (SEQ ID NO: 259)
FTHRNFVPTTST (SEQ ID NO: 260)
FTHRSSVPTTST (SEQ ID NO: 261)
FTHQSSVSTTST (SEQ ID NO: 262)
FTHQTSAPNTST (SEQ ID NO: 263)
FTHQTFAPNTST (SEQ ID NO: 264)
FTHQNSVPTTST (SEQ ID NO: 265)
FTHQSSMTTTRT (SEQ ID NO: 266)
FTHWIPVPTSST (SEQ ID NO: 267)
FTHWSPIPTTST (SEQ ID NO: 268)
FTHWSSGLTTST (SEQ ID NO: 269)
FHRSSVPTTST (SEQ ID NO: 270)
FNRSSVPTTST (SEQ ID NO: 271)
FNPWSSVPTTST (SEQ ID NO: 272)
FTQRSSVPTTST (SEQ ID NO: 273)
FTQRSSVPTTST (SEQ ID NO: 274)
FTQRSSVPTTST (SEQ ID NO: 275)
YNEPGLDEPPTT (SEQ ID NO: 276)
YAPQNLISIRGEY (SEQ ID NO: 277)

Exon 5

136 156
PGTSTVDVGTSGTPSSSPSPT (SEQ ID NO: 278)
PGTSTVDLRTSGTPSSLSSPTIM (SEQ ID NO: 279)
PGTSTVDLGTSGTPFSLPSPA (SEQ ID NO: 280)
PGTSTVDLG.SGTPSSLPSPT (SEQ ID NO: 281)
PGTSTVDLG.SGTPSLPSSPT (SEQ ID NO: 282)
PGTSTVDLGTSGTPSSLPSPT (SEQ ID NO: 283)
PGTPTVDLGTSGTPVSKPGPS (SEQ ID NO: 284)
PWTSTVDLGTSGTPSPVPSPT (SEQ ID NO: 285)
PGTSTVYWATTGTPSSFPGHT (SEQ ID NO: 286)
PGTSTVHLATSGTPSSLPGHT (SEQ ID NO: 287)
PGTSTVHLATSGTPSPLPGHT (SEQ ID NO: 288)
PDTSTMHLATSRTPASLSGPT (SEQ ID NO: 289)
PGTSAVHLETSGTPASLPGHT (SEQ ID NO: 290)
PGTSAVHLETTGTPSSFPGHT (SEQ ID NO: 291)
PGTSTVHLGTSETPSSLPRPI (SEQ ID NO: 292)
PGTSTIVNLGTSGIPPSLPETT (SEQ ID NO: 293)
PGTFTVQPETSETPSSLPGPT (SEQ ID NO: 294)
PGTPTVDLGTSGTPVSKPGPS (SEQ ID NO: 295)
PGTPTVYLGASKTPASIFGPS (SEQ ID NO: 296)
PKPATTFLLPPLSEATT..... (SEQ ID NO: 297)
QINFHIVNWNLSNPDPTSSEY (SEQ ID NO: 298)

Figure 7C

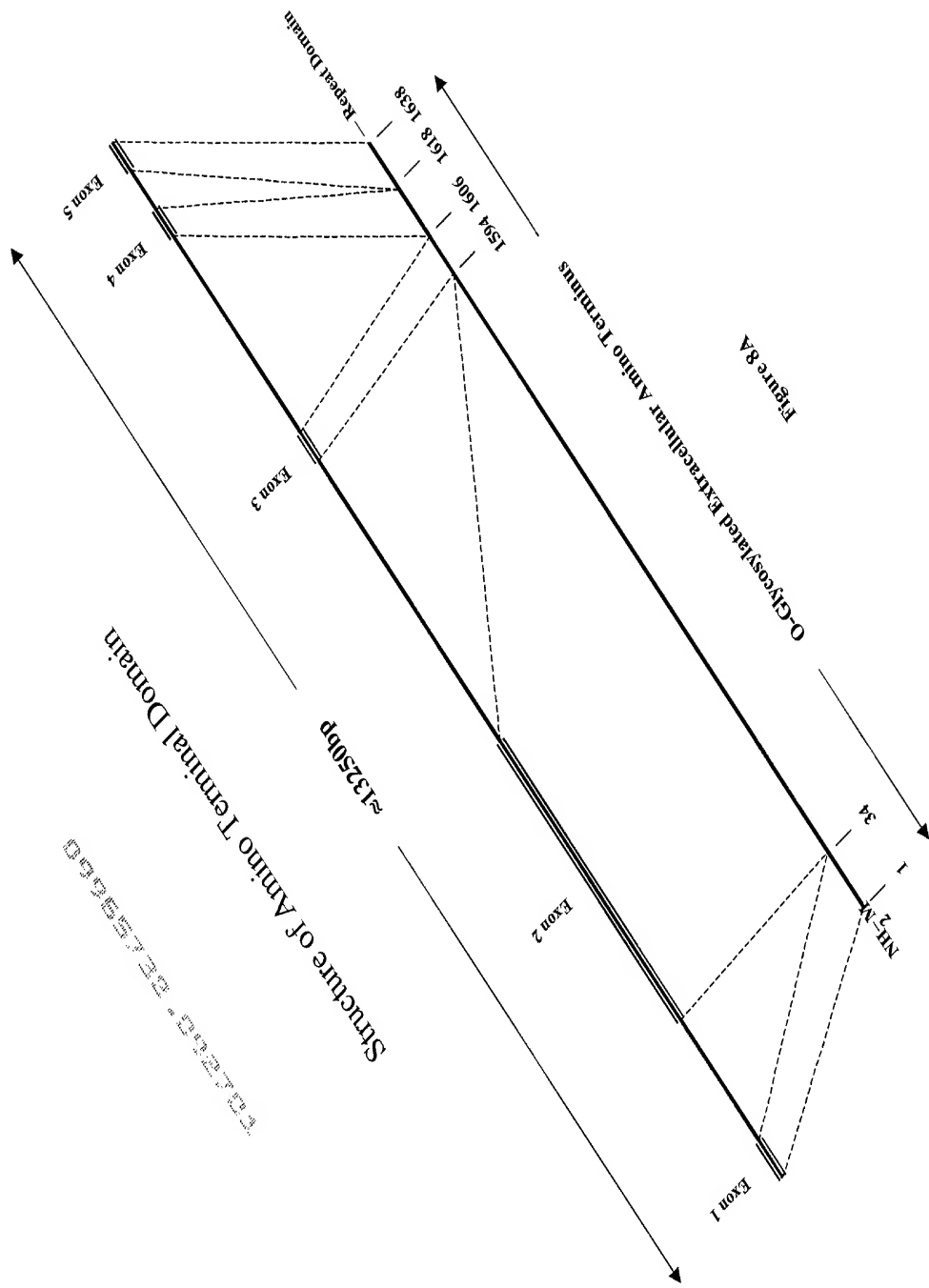


Figure 8A

7

Figure 8B (SEQ ID NO: 299)

NC_022860.1:362536..362566

Structure of Carboxy Terminal Domain

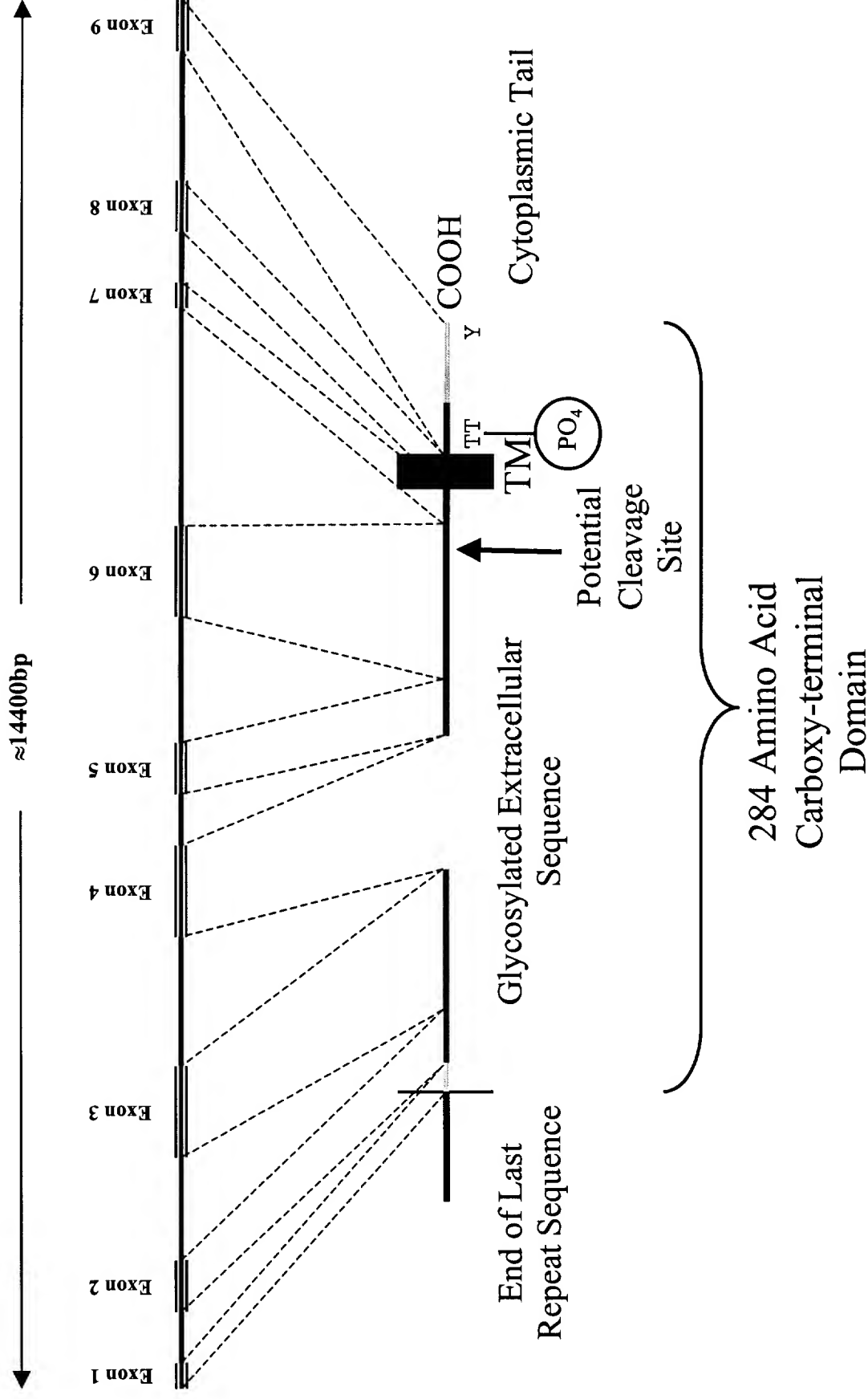


Figure 9A

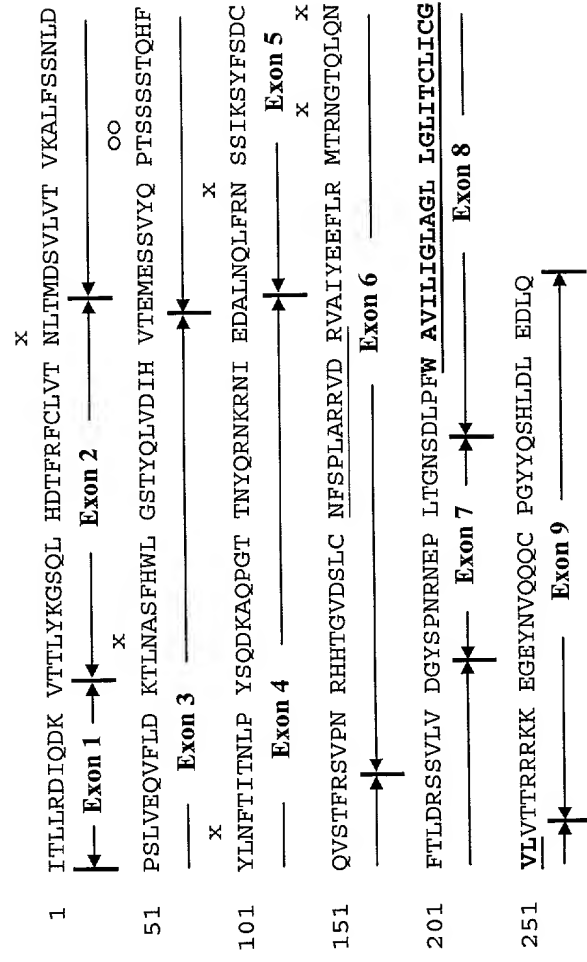


Figure 9B (SEQ ID NO: 300)

Proposed Structure of CA125

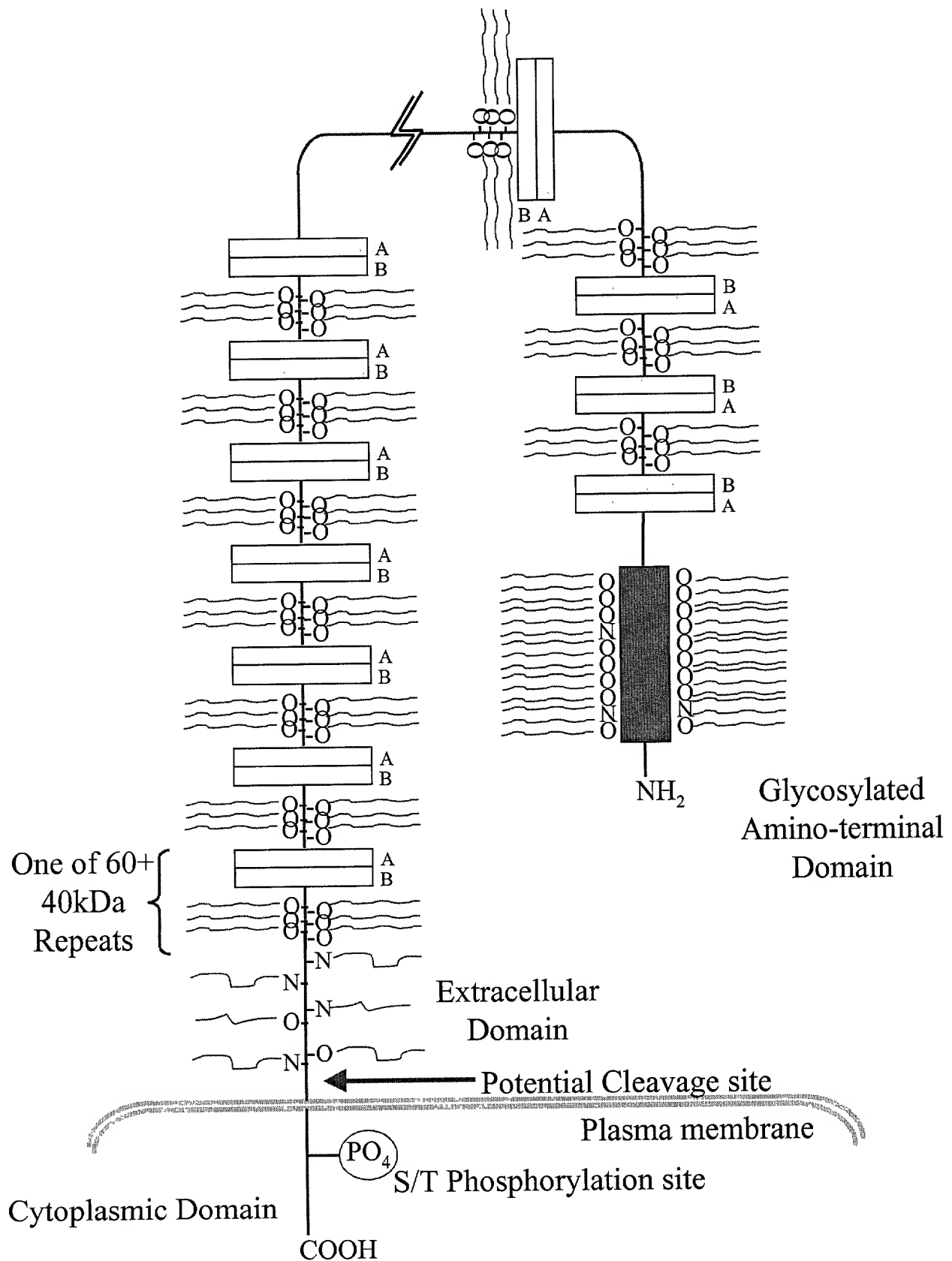


Figure 10